

**Desarrollo de una medida para la polarización
en las redes sociales.**

Bantú Agustín Valencia Sinisterra

Director: Dr. Frank Valencia

Resumen

Tras los últimos años el uso masivo de internet y las redes sociales han hecho que la polarización aumente, esto debido a las múltiples conexiones e influencias que se generan ahora, una sola persona puede ser influenciada por cientos de miles de personas en internet. Consecuencias de que la polarización aumente hace que arda sobretodo en temas, delicados conllevando a estallidos sociales como por ejemplo las protestas. Una alta polarización puede implicar la imposibilidad de que se llegue a un consenso, por eso es interesante tener una medida la cual te permita tener un índice de la polarización. Esteban y Ray desarrollaron una medida, esta presenta problemas, el principal es que no es suave ni continua, por el contrario, las ideas varían de forma suave, por lo tanto, lo más natural es que la polarización sea suave. Para ello vamos a utilizar unos modelos basados en DeGroot los cuales serán diferenciables, suaves y continuos. Adicionalmente con los modelos continuos se va a estudiar algunos casos donde la polarización va a tender a un mínimo usando la noción de consenso.

Palabras Clave: Continuidad, suavidad, polarización, influencia, consenso.

Introducción

Cómo se mencionó en el anteproyecto, la polarización es un fenómeno político y social el cual las personas se posicionan de forma opuesta a las demás respecto a un tema o varios temas en particular, formando entre dos o más zonas conglomeradas, este fenómeno está relacionado con los problemas sociales por ejemplo que cualquier idea moderada o intermedia a estas zonas pierda influencia o que es muy difícil encontrar un consenso entre dichas zonas. Además que en los últimos años la polarización ha aumentado, esto es debido al crecimiento de las redes sociales porque la comunicación entre los distintos individuos o agentes logran que las ideas que antes eran aisladas o que no tenían mucha repercusión lleguen a mucha más gente, esto sumado a la masificación de las noticias falsas, hacen que las redes sociales tiendan a ser extremadamente polarizadas, se puede ver claramente en las opiniones, de los republicanos y los demócratas o en Europa con temas como la inmigración, la independencia de Cataluña; en el caso de Colombia el proceso de paz, el paro nacional y las vacunas.

Fundamentación teórica

A continuación, se van a enumerar los temas y áreas que se van a tocar en este proyecto.

1. Cálculo diferencial. El estudio de la variación de las funciones de una variable
2. Cálculo integral. El estudio de las áreas, antiderivadas, volúmenes y series en funciones de una variable
3. Análisis matemático. El estudio de la variación, volúmenes, campos y funciones de múltiples variables.
4. Ecuaciones diferenciales. El estudio de las ecuaciones que relacionan la variable con su razón de cambio

5. Álgebra lineal. Área que estudia concepto como vectores, matrices, rectas y planos
6. Estadística. Área que estudia la variabilidad, interpretación y representación de un conjunto de datos.
7. Métodos numéricos. Rama que busca por medio de algoritmos encontrar soluciones que aproximen problemas de análisis matemático.
8. Computación paralela. Área que busca la creación de algoritmos que solucionan simultáneamente al dividirlo en subproblemas.
9. Sistemas dinámicos. Sistemas los cuales sus propiedades evolucionan en el tiempo.

Polarización de Esteban y Ray.

En [1] Esteban y Ray desarrollaron una medida de polarización basada en los grupos, primero particiona el rango de 0 a 1 de creencias en una cantidad de subintervalos que agrupan una cantidad de individuos con creencias dentro de cierto rango consideradas similares, creando un histograma además se toma un representante por cada grupo y con esto realizar la medida, esto se deberá hacer en cada instante de tiempo.

Este es un ejemplo del histograma para 4 grupos,

Vemos el anterior histograma definimos π tendrá las frecuencias relativas tal que

$$\pi_0 = \frac{2}{6}, \pi_1 = \frac{1}{6}, \pi_2 = \frac{1}{6}, \pi_3 = \frac{2}{6}$$

π_i es la proporción de agentes con estados de creencias en el grupo i . y_i valor representativo del grupo i , en este caso el valor medio del grupo i . Cada grupo i es un intervalo continuo $[a_i, b_i)$ que particiona el intervalo de 0 a 1.

Adicionalmente para cada grupo es necesario definir un representante, para esto usamos la variable $y \in \{0, \dots, m-1\} \rightarrow \mathbb{R}$, m es la cantidad de grupos

Para el histograma tenemos que

$$y_0 = 0.125, y_1 = 0.375, y_2 = 0.625, y_3 = 0.875$$

Ya que y tiene el punto medio en cada grupo del histograma.

Ya con esto podemos definir la medida de Esteban y Ray.

La fórmula de la medida de Esteban y Ray [1] es la siguiente:

Definición 11 medida discreta de Esteban y Ray

$$per(\pi, y) = K \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \pi_i^{1+\alpha} \pi_j |y_i - y_j|$$

Ecuación 1, medida de Esteban and Ray

K es una constante. α es un parametro. k es la cantidad de grupos.

Podemos verlo gráficamente con este ejemplo.

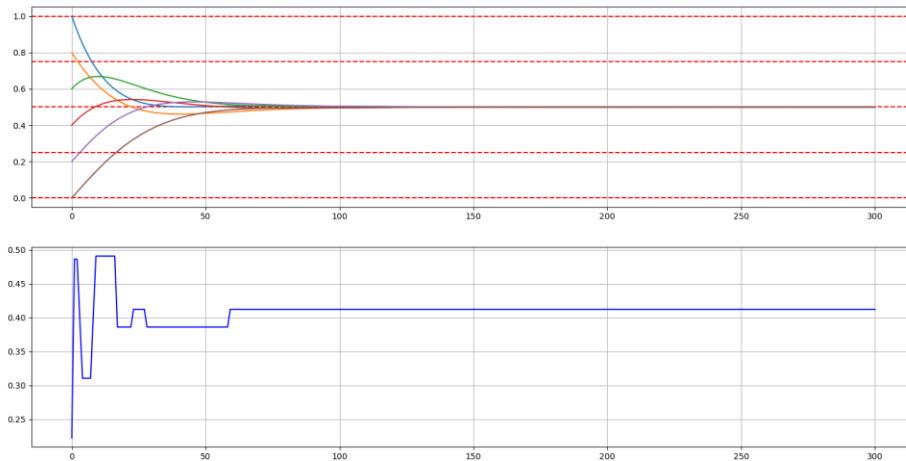


Figura simulación de la figura 1.

Vemos que cuando los grupos están alejados entonces la polarización es máxima cuando los grupos se acercan ligeramente, pero, aunque todos tiendan a 0.5, algunos agentes tienden a 0.5^+ y otros a 0.5^- por lo que, aunque vayan asintóticamente al mismo valor, no están nunca en el mismo grupo, por lo tanto, aunque todos van al mismo valor la polarización no es 0.

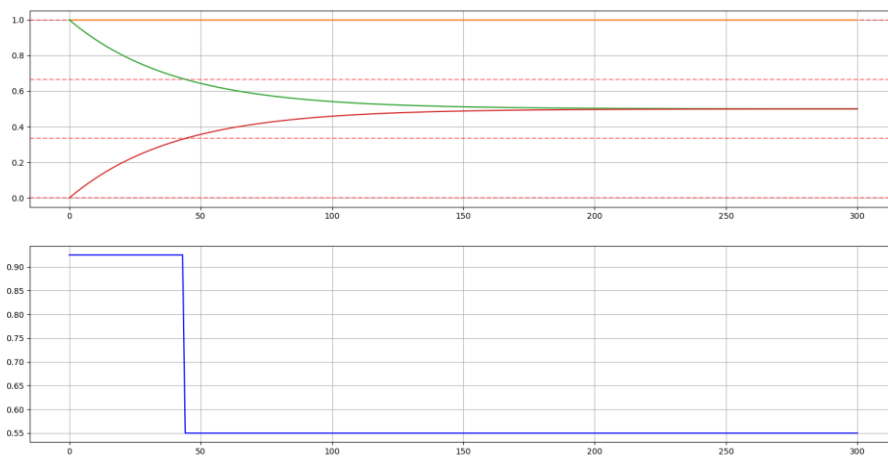


Figura simulación de la figura 2.

Vemos que cuando el agente verde y el agente rojo están en el mismo grupo la polarización toma una discontinuidad esto debe a la noción de los grupos

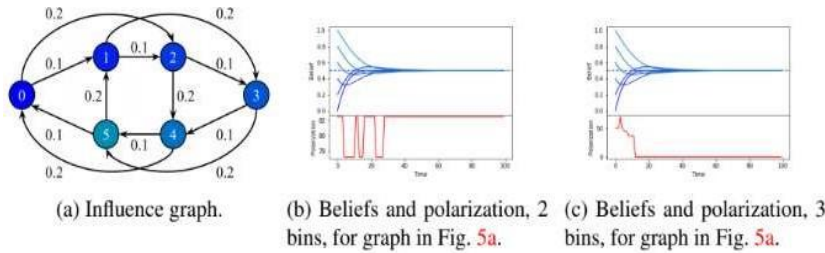


Figura 3. Diagrama que ilustra problema de discontinuidad. Tomada del Artículo de multiagentes [3].

En el gráfico podemos ver las relaciones de influencia que poseen cada agente con los demás en un grafo, en los otros dos vemos la evolución de los estados decreencia de cada uno de los agentes en el tiempo y la medida de la polarización, en el segundo caso con dos grupos y en el tercer caso con tres grupos. Esto se debe a que el estado de creencia de algunos agentes en el caso de los dos grupos tiende al límite a 0.5^+ y otros tienden a 0.5^- al esto ocurrir nunca pertenecerán al mismo grupo $[0, 0.5)$, $[0.5, 1)$. Aunque se acercan al mismo punto asintóticamente por eso la polarización en el caso de 2 grupos tiende a ser máxima. Se puede ver que la polarización en el caso de tener 3 grupos ($[0, 0.33\dots)$, $[0.33\dots, 0.66\dots)$, $[0.66\dots, 1)$) tiende a 0 eso se debe a que todos los estados de creencia de los agentes tienden a 0.5 como en el caso anterior, pero en este caso no hablamos con un borde de los grupos, al pertenecer al mismo grupo la polarización decrece y va directamente a cero, aunque los estados de creencia evolucionen de la misma forma la conclusión es distinta dependiendo de los grupos que se tomen, esto puede cambiar drásticamente las conclusiones que se tomen.

Resultados

Definición 1 (Definición polarización continua). Sea $P \in \mathbb{R}^n \rightarrow \mathbb{R}$ una función de vectores de tamaño n a los reales α un parámetro, f un KDE para $w(t)$, P es la medida de polarización del vector de creencias w al instante t y está definida por.

$$P(w(t)) = \int_0^1 \int_0^1 f(x)^{1+\alpha} f(y) |x - y| dx dy$$

α es una constante.

Ecuación 2, extensión continua de la polarización

A diferencia de la versión discreta, esta no requiere una noción de grupos.

Para demostrar que P es diferenciable y continua es necesario demostrar que P es derivable respecto a t .

Tenemos que:

$$\frac{dP}{dt} = \int_0^1 \int_0^1 (f(x)^{1+\alpha} f(y))' |x - y| dx dy$$

$$g(x, y; t) = f(x)^{1+\alpha} f(y)$$

Ecuación 3, derivada de la función de polarización

Por lo tanto, para que la derivada de P en el tiempo exista, la derivada en el tiempo de g debe existir

Aplicando la regla de la multiplicación y la regla de la cadena tenemos que

$$\frac{dg(x, y)}{dt} = f(x)^{1+\alpha} \frac{df(y)}{dt} + (1 + \alpha) * f(x)^\alpha \frac{df(x)}{dt} f(y)$$

Ecuación 4, desarrollo de la derivada de g

Note que si $w(t)$ es total y derivable, $f(x), f(y), \frac{df(y)}{dt}$ y $\frac{df(x)}{dt}$ existen, por lo tanto $\frac{dg(x,y)}{dt}$ existe, finalmente $\frac{dP(w(t))}{dt}$ existe. Es decir $P(w(t))$ es diferenciable y por lo tanto continua.

Para asegurar la continuidad de la función f debemos tener un modelo de creencias basado en DeGroot continuo, por lo que tenemos el siguiente.

Para el sesgo de autoridad tenemos la siguiente ecuación:

$$w'_i(t) = \frac{1}{n} \sum_j^n I_{i,j} * (w_j(t) - w_i(t))$$

Ecuación 5, extensión continua para el sesgo de autoridad

Análogamente para el sesgo de confirmación

$$w'_i(t) = \frac{1}{n} \sum_j^n B_{i,j} * I_{i,j} * (w_j(t) - w_i(t))$$

Ecuación 6, extensión continua para el sesgo de confirmación

Con estas ecuaciones solo basta que w sea una función total en los reales para que w' exista en todos los reales y asegurar la continuidad de f y por lo tanto de P .

Conclusiones

A lo largo de este proyecto se han usado múltiples herramientas de cálculo y las ciencias de la computación para lograr obtener un método para medir la polarización de un grupo de individuos, donde esta presenta suavidad, continuidad y además cumple con los axiomas necesarios para poder capturar la polarización. Para ello usamos los kernel de densidad, las integrales múltiples y además se logró extender el modelo de multiagentes de DeGroot a un caso continuo para poder asegurar la continuidad de la función de polarización, dando soluciones analíticas para algunos casos específicos y gráficos respecto al tiempo.

Las contribuciones presentadas en esta tesis son de carácter teóricas y computacionales, en particular se puede observar cómo usando herramientas del Cálculo, Estadística y Algebra nos permite tener un sistema que sea capaz de desarrollar un modelo de polarización en el tiempo que esté definido para cualquier instante de tiempo, que cumpla los axiomas necesarios para capturar la polarización y que además sea continua y suave, lo que permite reducir errores dados por discretizar la polarización, incluso en algunos casos es posible tener una fórmula para obtener la polarización en cualquier instante de tiempo. Por lo tanto, se concluye que se pueden aplicar las áreas ya mencionadas para lograr resolver el problema de la polarización en particular.

Además, se puede también concluir que gracias a los avances obtenidos en esta tesis podemos darle mejores estimaciones a la polarización debido a que redujimos errores que no correspondían a la discretización de la polarización, lo que se tiene métricas para la polarización que no tiene problemas de la discretización.

Finalmente, para la aplicación de estas técnicas se usan métodos numéricos para estimar la polarización usando métodos para aproximar ecuaciones diferenciales y métodos para aproximar integrales, para esto se usa la aceleración por hardware y en este caso particular pycuda y numba, para usar los algoritmos de ecuaciones diferenciales usamos el método de Euler para ecuaciones diferenciales y el método de Simpson para poder tener buenas estimaciones de integrales y de ecuaciones diferenciales. Gracias a eso podemos estimar el valor a un punto específico y poder graficar para ver las tendencias de la polarización.

Bibliografía

- [1] Esteban and Debraj Ray. (1994). On the Measurement of Polarization
<https://www.jstor.org/stable/2951734?origin=JSTOR-pdf>
- [2] Esteban and Ray's. (2012). Comparing Polarization Measures.
<https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780195392777.001.0001/oxfordhb-9780195392777>
- [3] Mario S. Alvim, Bernardo Amorim, Sophia Knight, Santiago Quintero, and Frank Valencia Bias. (2021). Polarization and Belief Convergence of Multiagent Systems under Confirmation.
<https://arxiv.org/abs/2104.11538>